UNCLASSIFIED

AD-A202 292

DTIC FILE COPY

(2)

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

DTIC ELECTE DEC 0 9 1988 S D

D Cb

1b. RESTRICTIVE MARKINGS	
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
3. DISTRIBUTION / AVAILABILITY OF REPORT	Approved for public release; distribution unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	
5. MONITORING ORGANIZATION REPORT NUMBER(S)	AFOSR-TR- 88-1224
6a. NAME OF PERFORMING ORGANIZATION	Duke University
6b. OFFICE SYMBOL (If applicable)	
7a. NAME OF MONITORING ORGANIZATION	AFOSR/NL
6c. ADDRESS (City, State, and ZIP Code)	Durham, NC 27706
7b. ADDRESS (City, State, and ZIP Code)	Bldg. 410 Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	AFOSR
8b. OFFICE SYMBOL (If applicable)	NL
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	AFOSR-87-0353
8c. ADDRESS (City, State, and ZIP Code)	Bldg. 410 Bolling AFB, DC 20332-6448

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2313	A6	

11. TITLE (Include Security Classification)

Categorizing Sounds

12. PERSONAL AUTHOR(S)

Dr. Gregory R. Lockhead

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM 1 Sep 87 TO 30 Sep 88	30 Sep 88	56

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Much of the variability in judgments of univariate sounds depends on what stimuli occurred recently (sequence effects), what stimuli might occur (set and range effects). A model examined in this report associatees much of this response variability with two factors, assimilation in order to maintain a veridical response scale. Studies of univariate stimuli reported here show sequence effects that are consistent with the model and not consistent with an attention-band model. Context effects in multidimensional judgments were also examined. Identification of these multidimensional stimuli was superior to that of univariate stimuli, but responses again assimilated toward the value of the prior stimulus. Keywords: ... Acoustics, ... Intensity, Sequence Effects, ... Judgment ...

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	Unclass

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Dr. John F. Tangney	(202) 767-5021	NL

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

1

UNCLASSIFIED

CATEGORIZING SOUNDS

Gregory R. Lockhead

Department of Psychology

Duke University

Durham, North Carolina 27706

Grant No. AFOSR-87-0353

Interim Report

Period covered: 1 September 1987 to 30 September 1988

30 September 1988

Grant Monitor: John F. Tangney, Program Manager,

Life Sciences Directorate

Prepared for: Directorate of Life Sciences

Air Force Office of Scientific Research

Bolling Air Force Base

Washington, D. C. 20332-6448

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12.
Distribution is unlimited.
MATTHEW J. KERPER
Chief, Technical Information Division

Approved for public release;
distribution unlimited.

DTIC COPY INSPECTED 6

Accesion For	
NTIS CRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

13 OCT 1988

CONTENTS

Categorizing Sounds

Gregory R. Lockhead
Duke University

Abstract

What a sound or other stimulus is identified to be depends on context. Much of the variability in judgments of univariate sounds depends on what stimuli occurred recently (sequence effects), what stimuli might occur (set and range effects), and what other information is available to the subjects (task effects). A model examined in this report associates much of this response variability with two factors, assimilation in memory and how subjects adjust for assimilation in order to maintain a veridical response scale. Studies of univariate stimuli reported here show that responses are less variable when subjects have an identification function (IDF) for mapping auditory amplitudes onto responses than when they do not have an IDF, and show there are sequence effects that are consistent with the model and not consistent with an attention-band model. Context effects in multidimensional judgments were also examined. People were asked to identify sawtooth correlated tones that varied across trials in loudness and in pitch. Identification of these multidimensional stimuli was superior to that of univariate stimuli, but responses again assimilated toward the value of the prior stimulus. In a different study, people judged loudnesses when tones varied randomly in pitch. More time was required to classify these multidimensional stimuli than to

classify univariate tones. Further, response times were the longest in conditions having the largest pitch differences. This set of results is consistent with the memory model and reveals some previously unknown aspects of what is involved when people attempt to identify sounds.

On Categorizing Sounds

What a sound or any other stimulus is identified to be is not determined just by that stimulus. Judgments also depend on other stimuli that might occur. This project explores four facts concerning context: What a sound is identified to be depends on (1) the differences between all of the sounds that might occur during the study (range effects), (2) the order in which those sounds occur (sequence effects), (3) for multidimensional stimuli, the ways univariate attributes are combined to produce complex sounds (set effects), and (4) what information is given to the subject (task effects).

Each of these context effects - range, sequence, set, task - is important both theoretically and practically. Practically, range can modify responses by a factor of six or more, sequence and task differences can each affect choices by 75% or more of the total response range, and set can move identification accuracy from nearly chance to nearly perfect (cf. Lockhead, 1984). Theoretically, as Luce and Krumhansl (1988) have concluded, psychophysical models are and will remain difficult to evaluate unless response variance due to context is accounted for. Also, and independent of measurement issues, context effects provide information for testing models of memory and choice.

Background.

The general goal of this project is to better understand how

complex sounds are identified. The specific goals are to further evaluate a proposed model of sequence effects in univariate tasks and to learn if the model generalizes to predict multidimensional judgments.

Recent studies suggest it is useful to examine effects of context on average choices, on response variability, and on response times. The logic for one such studies was based on these observations of absolute identification data: Response variability is larger in stimulus sets having a larger stimulus range, sequence effects (particularly assimilation) are larger on trials that successive stimuli are more different, and, necessarily, successive stimuli are more different more often in conditions in which the stimulus set has a larger range. Combining these observations, Lockhead and Hinson (1986) showed that a large portion of the context effect on performance that is associated with stimulus range (a between-conditions measure) is due to stimulus sequence (a within-conditions measure). Although these results have been treated separately in the literature, they are apparently due to a common source. Whenever successive stimuli are more different, whether in a between-conditions or a within-conditions design, performance is less competent.

Lockhead and Hinson (1986) used college students as subjects and tones as stimuli. Hinson and Lockhead (1986) replicated that study , except with pigeons as the subjects and flickering lights as the stimuli. The purpose was to learn if mechanisms analogous to those responsible for context effects in humans are also available to simpler subjects.

They are. The pigeon data were essentially identical to the human data. Hinson and Lockhead (1987) then showed that both sets of results are consistent with a memory model in which judgments are partially based on the subject's biased memory of prior stimuli. According to that model (cf. Lockhead & King, 1983), the response to a stimulus, R_N, assimilates toward the value of the stimulus on the just previous trial and contrasts from values on earlier trials:

$$R_N = S_N + a(M_{N-1} - S_N) + b(\bar{M} - M_P) \qquad (1)$$

where S_N is the stimulus, M_{N-1} the memory of the previous stimulus (estimated in practice as S_{N-1}), $\bar{M}$ is the the average memory of all stimuli during the experiment, M_P is the average memory of stimuli on trials N-2 to N-7 and called the memory pool (estimated in practice by the mean stimulus on those trials), and a and b are positive constants. When feedback is not given, b is close to zero.

Equation (1) describes assimilation between successive responses and contrast between the response and the average of several prior stimuli. The assimilation aspect of the model is diagrammed in Figure 1. This formulation describes much of the response variability in several data sets (Lockhead, 1984).

---insert Figure 1 here---

Response variability.

The fact that there are sequence effects (SEs) in judgment

data means that the response to any particular stimulus depends in part on what happened on prior trials. Performance is not determined entirely by the stimulus being judged.

The most prominent SE in judgment data is assimilation. It is regularly found that responses are overly similar to the value of the prior stimulus (or response, depending on details of the experiment). Stevens (1975) called this effect hysteresis and attributed it to response bias and stimulus order. Because his primary interest was in an underlying psychophysical function, Stevens suggested averaging data across sequences in order to average out context effects and thus better reveal the power law. The procedure is reasonably successful. In log-log coordinates, averaged magnitude estimations (MEs) are often linear with stimulus intensity. Nine examples are shown in Figure 2.

Response variability was not reported with the data in Figure 2. This is a common practice. This variability is not directly relevant to the power law. Too, the standard error of the judgments can be made arbitrarily small given enough estimates to average across, and so any contribution of response variability can effectively be eliminated (Stevens, 1975, p. 440).

---insert Figure 2 here---

Stevens' power law describes <u>average</u> responses. One goal of the current project is to predict <u>momentary</u> responses, the response on any given trial. Because each response depends on prior trials, average responses are not sufficient for this

purpose. In order to predict responses on individual trials, it is important to account for response variability associated with SEs.

It is possible that this is only an academic concern. The response variability associated with context might be so small as to be of little importance and thus might be ignored. The present study examined if this is the case.

Method and Procedure. Two people were asked to give MEs to 30 randomly presented tones (0.5 sec duration, 1000z. sine waves) that ranged from 51 to 80 dBA, in 1 dBA steps, with a modulus of 100 assigned to the 65 dBA tone. Each of two subjects gave 2,000 responses.

Results. The average response to each stimulus by each subject is shown in Figures 3a and 3b. Responses increase monotonically with intensity and show little other variability. These results are generally consistent with the power law. While the functions are not perfectly linear, this is probably only because subjects' use different numbers differently (Baird & Noma, 1975; Teghtsoonian & Teghtsoonian, 1988).

---insert Figure 3---

The individual responses that were averaged to produce Figures 3a and 3B are shown in Figures 4a and 4b. These data are quite variable. For example, subject EM called the 51 dBA stimulus "5" on 21 of the 87 times it was presented, "10" 36 times, "20" 17 times, "30" 5 times, "40" 4 times, and "50" 4 times; a ten-fold response range. Both subjects assigned many

different responses to each stimulus. Typically, the range of responses to any stimulus was about half the total response range.

---insert Figure 4---

Much of this response variability is not simply random. Responses tend to be small on trials that the prior stimulus was small, and large when the prior stimulus was large. This contingency between the response and prior stimulus is seen in Figure 5. The geometric mean response to all stimuli was larger on trials that the stimulus on the prior trial was larger, and smaller when the prior stimulus was smaller. This is assimilation. Assimilation has been reported in every ME and absolute identification (AI) data set that have been examined and reported in the literature.

---insert Figure 5---

Averaged MEs are commonly reported in terms of Stevens' power law:

$$R = kI^b \qquad (2)$$

where R is the mean magnitude estimation, I is stimulus intensity, and k and b are constants. For the task of predicting responses on any particular trial, while maintaining the spirit of the power law, the fact of assimilation suggests modifying Equation (2) in terms of Equation (1):

$$R_N = I_N{}^b + \underline{a}\,(I_N - I_{N-1})^b + b(I_M - I_P) \qquad (3).$$

Equations (1) and (3) predict individual judgments _very_ much better than does Equation 2 (Lockhead & King, 1983).

A Decision Theory View.

One interpretation of assimilation is that the position of the criterion for judgment depends on the prior trial. When the prior stimulus was large the criterion shifts lower, and when the prior stimulus was small the criterion shifts higher (Holland & Lockhead, 1968; Triesman & Williams, 1984). According to this view, the position of the criterion depends on sequence.

If this is the case, then d′ measured according to statistical decision theory (SDT) and calculated from averaged data will be smaller (indicating poorer performance) than d′ calculated from individual stimulus sequences. This is because the varying criterion adds noise to the overall data, and because some of this noise is removed when sequence is taken into account.

In order to interpret data in this manner, it is necessary to reformulate one aspect of the traditional SDT analyses. In ordinary SDT accounts, the criterion is assumed to be fixed during a set of trials, and any noise is assumed to be associated with the system. The reformulation suggested here is, instead, to consider the system as fixed over a set of trials and to assume the noise is associated with the criterion.

This formulation does not pose any difficulty for statistical decision theory. It is mathematically equivalent to all features of that theory to assume the noise is in the criterion as to assume it is in the signal (Green & Swets, 1974). The traditional assignment of noise to the system rather than to the criterion is an arbitrary convention.

The advantage of associating noise with the criterion is that variability associated with sequence can then be accounted for, rather than being treated as random error.

This approach may have some merit. Lockhead and Hinson (1986) calculated d′ separately for different stimulus sequences in data from a three-stimulus absolute identification study. They also calculated d′ based on averaged data. The d′ measures were larger when variability due to criterion changes was taken into account. This again indicates that performance in judgment depends lawfully and measurably on what occurred on the prior trial.

Memory of the stimulus.

The above discussions show that stimulus judgments are variable, that criterion positions depend on sequence, and that response selections depend on inferred possibilities. These facts help describe the essential thesis of this project: Each stimulus is compared to the memory of the prior stimulus, this memory is biased by assimilation, and responses are selected from the set

of known possibilities in terms of this representation.

It is not known if such a memory interpretation is correct. While SEs might be associated with memory, there also are other possibilities. SEs might, instead, be due to sensory mechanisms, or encoding, or perceptual or response processes, or some combination of these. All of these have been suggested in the literature.

If we assume that SEs are due to one just process, then we can reject both stimulus intensity and sensory response as the cause. The reason is there are SEs when there are no intensities, and thus when there also is no sensory response. For example, if the signal generator is turned off and people are asked to guess what might have been presented even though there was no actual stimulus, there are again SEs. Moreover, these SEs are identical in form to the SEs observed when actual intensities are classified (Ward & Lockhead, 1970). Since the same effects occur whether or not there are stimuli, the effects cannot be due to just stimuli.

SEs are also not due just to biased memories of prior intensities. This is again for the previous reason. There are no intensities to be remembered, biased or not, in guessing studies in which the stimulus generator is turned off, but there are still SEs.

Of the alternatives suggested to date, this leaves the single possibility that SEs are due to response processes, which might additionally be modified by feedback.

We have previously suggested this is the case, that SEs are due to response system processes (Lockhead & King, 1983). While we still consider this to be a well supported inference, we further consider here that the situation may be more complicated. Unfortunately, SEs might not be due to a single factor. Rather, they may be associated with whatever <u>information</u> the subjects have available to use in performing the assigned task, and this can be different in different situations. That is, while response processes may always be involved, memorial and other processes might also involved in many judgment situations.

This inelegant interpretation, that the source of SEs in not a single factor, is at least consistent with the many complexities in judgment data. SEs depend on the duration between stimuli, the presence or absence of feedback, the form of the feedback, whether or not the stimulus generator is functioning, the physical difference between successive stimuli, whether the stimulus modality varies between trials, and the distribution of stimuli in the set. <u>There are always sequence effects</u>, and their form and magnitude depend on details of the situation (Lockhead, 1984).

The following discussion further examines this suggestion that SEs are associated with whatever is available for the subjects to use in order to maintain a reliable response scale.

Identification Functions.

One feature of the information important to subjects in judgment tasks is whether or not there is an identification function (IDF), and if that IDF is known to the subjects. A known IDF exists whenever subjects have been shown a fixed stimulus set. An example is absolute identification (AI) studies with feedback in which the subjects know there are N unique stimuli and in which the appropriate mapping of each is told the subjects on each trail.

There is not always a known IDF, whether or not an identification formally exists. Magnitude estimation (ME) studies are an example. There is a formal IDF in ME tasks because the experimenter uses N discrete stimuli. However, subjects are not told this. Rather, the subjects are allowed to believe that any of perhaps thousands of stimuli are possible on any trial. If assimilation affects the perception of each stimulus (or the memory of prior stimuli), the subjects may never realize there is an IDF and that only a few different stimuli are involved in the study. Rather, they may perceive many more stimuli than actually exist.

There is a difference in AI and ME data that is consistent with this observation. Response variability is non-monotonic with intensity in AI loudness judgments (IDF known to the subjects) but is monotonic with intensity in ME loudness judgments (Lockhead & King, 1986, 1988).

This difference in performance may be related to inferences made by the subjects concerning stimulus and response alternatives. In AI tasks, the subjects know that certain stimuli are not possible. For example, if the last stimulus was judged to be #2, and if and the current stimulus is judged to be smaller than that one, then only response #1 is available and the subject will report #1.

The situation is different for the same sequence in an ME task. Here there is no restriction on the response. This is because the subject is not aware of any restriction on the stimulus possibilities. Thus, if assimilation affects perception or memory as appears to be the case, many different responses might be made.

Said differently, in an ME task the subject reports the perception of the stimulus, while in an AI task the subject judges which of the N possibilities was presented. If this interpretation is correct, then there should be many different features in data sets for which there was a known IDF than in data sets for which there was no IDF. The possible difference explored ahead is the variability of responses to individual stimuli.

The data from the ME study using 30 stimuli reported in Figure 4 are highly variable. The variability in the data of the nine studies in Figure 2 is not known. Only 5 to 7 stimuli were involved in those. Had those stimuli been used in absolute identification (AI) tasks, there would be little response variance. AI performance is very good when there are few, widely

spaced stimuli.

The ME data might also have had little variability. If so, then the marked variability in Figure 4 might be because there were many stimuli (30) or the subjects were inattentive. On the other hand, if Stevens' data were also highly variable, there would be an implication concerning the source of SEs and thus the resulting variability. There is a known IDF in AI tasks but not in ME tasks, and this becomes a candidate account for noise that occurs in ME but not AI data sets.

More specifically, in an ME task, each stimulus assimilates in memory toward prior memories, the subject has no information of this distortion, and the subject also has no knowledge that only 6 rather than thousands of different stimuli are possible. Therefore, because of assimilation, the subject has memories of many, many different stimuli in the study. Thus, many different responses will be assigned to the same stimulus on different trials.

The number of these possibilities in memory, and thus response possibilities, could be large. Each stimulus is preceded by, and thus assimilates toward, every other stimulus. If there are 6 stimuli, this produces 6^2 memories. If the magnitude of assimilation also depends on the physical difference between successive stimuli, then there would be even more memories. By this analysis, the same physical stimulus should be assigned many different responses during the experiment.

The situation is different when there is a known IDF. In AI

tasks, the subjects know how many stimuli are to be categorized. Thus, they should respond to each stimulus with its best fitting category, not just with what the intensity appears to be. If the magnitude of assimilation is less than a half category step, category performance could then be perfect in such AI tasks, even though perception is not perfect.

A study. To estimate the amount of response variability in Stevens' data, we asked four people to give 300 MEs of the loudnesses of six tones. These were 1200 Hz sinewaves separated by 4 dB steps. As a control, we asked another two people to give 300 absolute judgments of the same stimuli.

The AI performances were nearly perfect. The ME performances were highly variable. Apparently, what people judge in ME tasks is not just the stimulus. Rather, they compare its biased appearance with many biased memories.

There is secondary support of this conclusion. Once the study was completed, the subjects in the ME task were asked to estimate the number of different stimuli heard during the study. Their estimates ranged from 80 to 100 tones. Implications of this are currently being pursued in studies that relate performance to the subjects' knowledge of the stimulus set.

The neural attention-band and assimilation in memory. Luce, Green, and Weber (1976) and Luce and Green (1978) have proposed that SEs in ME data are due to a peripheral neural process. Their suggestion is based on two findings: Successive responses

are positively correlated when successive stimuli are similar, and the magnitudes of these correlations decrease with increases in the difference between successive stimuli. These results were taken as support of an auditory attention-band model.

This attention-band is equated with a differential ability of the system to sample incoming neural fibers. Aspects of two signals falling within the hypothesized band are less variably estimated than are those same aspects when they fall outside the neural band. When successive signals are within the band, observers are presumed able to preserve the ratio of the successive intensities in their responses, and the responses will be correlated. If one of the signals falls outside the band, as when the successive stimuli are very different, the sample is then not sufficient to allow the ratio of the intensities to be incorporated in the response, eliminating the correlation. Separation in dB is involved in the predictions because the attention band is assumed to be narrow and centered on the past stimulus intensity.

The attractions of this attention-band hypothesis are several. A neural mechanism is implicated, the idea is easily communicated, the theory describes a complex data set, and the term "attention" suggests relations between psychological concepts and a neural structure.

There are also reasons not to be attracted to this hypothesis. First, an intellectual leap is needed to relate verbal responses to a specific neural system. Second, successive responses are also correlated in guessing tasks although there

are then no intensities to be involved in a neural band. Third, the memory model also predicts these and other results not addressed by the attention-band model. Fourth, as shown ahead, the noted correlations are a function of the IDF and not intensity to the ear.

Method and Procedure. We collected MEs with and without feedback given after each response. The tasks were magnitude-estimation (ME) and magnitude-estimation-with-feedback (FB). The stimuli were 30, 1,000 Hz sinewaves differing in 1 dB steps from 51 to 80 dB.

The ME task was conducted first, with the 65 dB intensity assigned modulus 100. The best fitting power function was calculated for those averaged data. That equation provided the (rounded) feedback numbers used in the FB task.

For tasks, two observers produced 2000 responses in five experimental sessions.

Results: Magnitude Estimation. Figure 6 shows the correlations between successive responses as a function of the physical difference between successive stimuli. The correlations are positive and large when the differences are small, and the correlations decrease in magnitude as successive stimuli become more different. This replicates several earlier reports by Luce and colleagues.

---insert Figure 6---

However, these correlations do not simply drop from positive to zero as successive stimuli become more different. They become increasingly negative (to perhaps $r = -0.7$) when the stimuli are very different.

Negative correlations are also seen in some published graphs, although they have not been commented upon. However, most published data show only positive correlations. This may be an artifact of the reporting method. Figure 6 shows the correlation for <u>each</u> stimulus difference. The practice in the literature has been to collapse the data over groupings of a few stimulus differences for similar tones and many stimulus differences for dissimilar tones. The probable reason is there are fewer occasions of large than small stimulus differences. For example, using 30 stimuli in 1 dB steps, 30 occasions of a 0 dB difference (stimulus repetition) are expected for every one occasion of a +29 dB difference. While collapsing the data differently for different differences partially adjusts for this discrepancy, it also masks effects that might exist at large stimulus differences.

---insert Figure 7---

A different analysis that increases the number of data points while retaining a measure for each difference is to combine the data over absolute differences between successive stimuli before calculating the correlations. When this is done for the current data, all nine correlations for the largest differences between successive tones, those of more than 20 dB, are negative.

While the attention-band model predicts the positive correlations seen in Figures 6, it does not account for the negative ones. As described in the next section, the memory model attributes the positive correlations to assimilation, and attributes the negative correlations to adjustments made from time to time by the subjects to correct for the constricted response scale.

Feedback. In ME, each response assimilates toward the prior response. If this were the only biasing process involved in the judgment task, then continuing response compression would eventually result in every stimulus being assigned the same response. This does not happen. Apparently, this is because the response compression becomes noticed, particularly when stimulus differences are large. When that occurs, the subjects attempt to correct for the discrepancy by expanding the response range. This would produce negatively correlated successive responses when successive stimuli are very different.

This interpretation predicts that all of the correlations between successive responses in ME data should become smaller or disappear when the task is changed, so that assimilation now occurs to something other than the prior response. In AI experiments with feedback, it is known that assimilation is between the response and the prior feedback, rather than the response and the prior response. This is apparently because the feedback is the subjects best estimate of the correct response. If this is the case, then there should also be no or little correlation between successive responses in ME when feedback is

given. There should be no positive correlations because assimilation would be to the feedback rather than the response, and there should be no negative correlations because there are no cumulative effects on the response scale to be corrected.

The correlations for the feedback data are shown in Figure 7. The large positive correlations seen in ME data (Figure 6) are not seen here. Only 38 of the 60 correlations are positive, and even these are not associated with small differences. The largest ones (at -29, -1, 11 and 25 dB differences) are scattered rather than clustered at small differences. Further, none of these correlations approach the frequent magnitude of +0.7 seem in the feedback data (Figure 7).

---insert Figure 7---

There is also no trend to negative correlations at large stimulus differences. Rather, the correlations are erratic across stimulus differences. Since the number of observations entering into these calculations is small, these are the results expected if there is little or no real effect. Collapsing the data over positive and negative stimulus differences increases the stability of the data and reinforces the conclusion that there is no trend for large correlations.

Figure 7 also shows a general tendency for the correlations to be positive, rather than zero or negative. This is expected if assimilation is between the response and the prior feedback. This is because performance is above chance and so responses and stimuli are correlated.

According to the attention-band account, ME and FB data sets should both have the same positive correlations between successive responses when successive stimuli are similar. This is because the intensities in both tasks were the same. The correlations are not the same. According to the memory model, the correlations in the two data sets should be different, and they are. This is because feedback provides for response adjustments on each trial. Information for adjustment is not available when there is not a known IDF.

These data do not support the neural attention-band hypothesis. They do support the views that assimilation occurs in memory, not at the ear, and that subjects use whatever information is available to adjust the resulting distortions of the response scale.

Context effects in multidimensional judgments.

Absolute identifications (AI) of univariate stimuli are poor (slow with many errors) when people judge a large number of stimuli. An example used in this report is that people make many errors when they attempt to uniquely identify 10 sinewaves that differ in 2 dB intensity steps or that differ in 5 Hz frequency steps.

Identification performance often improves when multidimensional stimuli are judged. For example, if the 10 loudness and pitches are combined redundantly and linearly, such

that successive stimuli differ from one another in 2 dB <u>and</u> 5 Hz steps (cf. the open dots in Figure 8), identification of these bivariate stimuli is slightly better than that of either univariate set. And if these univariate sets are combined redundantly but now nonlinearly (cf. the filled dots in Figure 8; this particular redundant configuration is known as sawtooth-pairing) identification performance improves much more. In some reports, performance has been about 30% correct for univariate stimuli (any row or column in Figure 8), 35% correct for linearly correlated stimuli (the major or minor diagonals), and nearly perfect for sawtooth-paired stimuli (Lockhead, 1970).

---insert Figure 8---

Context effects have not previously been examined in such multidimensional data. The present study measured sequential structure in data when sawtooth-paired stimuli were judged in an AI procedure.

<u>Method</u> <u>and</u> <u>Procedure</u>. The stimuli were ten tones as indicated by the filled circles in Figure 8. There were 10 levels of loudness (79 to 88 dB SPL in 1 steps) and ten levels of pitch (1000 to 1045 Hz in 5 Hz steps), with each amplitude paired with one and only one frequency. This provided ten redundant stimuli.

Using the response categories 1 - 10, people were asked to identify the intensity of each tone when feedback (the numerals 1 - 10 corresponding to the intensity level) was given (feedback conditions) and was not given (no feedback conditions) after each response. Four different subjects gave 300 responses in each of

ten sessions, for 3,000 responses per subject per condition.

For this study, the subjects were informed about the structure of the intensity-frequency correlation. Figure 8 was mounted above the response keyboard and could be referred to whenever useful to aid the subjects in judging which loudness was presented. Thus, there was a known IDF. The subjects responded by pressing one of ten keys on a special 10-key keyboard in which the a finger fit comfortably over each key. They were asked to perform as quickly but as accurately as possible.

Since errors are needed to measure sequence effects on judgments, the stimulus values in this study were selected during pilot work to be difficult to identify accurately. Overall, mean accuracy in identifying these sawtooth paired stimuli was 46% correct when feedback was given and 41% correct when there was no feedback.

Results: Feedback Given. The mean response to each stimulus as a function of the prior stimulus was averaged over subjects and sessions. These results are shown in Figure 9, with amplitude along the abscissa and frequency along the ordinate. Filled dots indicate the 10 stimuli (cf. Figure 8). The 10 irregularly outlined sets of numerals indicate the mean response to each stimulus (the nearest filled dot indicates its related response set) when the prior stimulus was the numeral listed within the outline. [There is no meaning of the encircling except to indicate what mean response is associated with what current stimulus and what prior stimulus.]

---insert Figure 9---

For example, consider the lower-left response set in Figure 9. This shows the mean responses to stimulus #1 (amplitude 1 or 79 dB; frequency 3 or 1010 Hz) in Euclidean coordinates of amplitude and frequency steps when the rior stimulus (the encircled numerals) was each of the ten possibilities. It is seen that the average response to stimulus #1 was near the physical position of stimulus #1 on those trials that the prior stimulus was also #1. The mean response to stimulus #1 was further from this value when the prior stimulus was larger than #1, and was farthest from it when the previous stimulus was the most distant from #1, when it was #10.

This is assimilation. The response tends toward the value of the prior stimulus. This sequence effect is numerically greatest for central stimuli (e.g., #5 and #6). There, judgments depend on the previous stimulus by as much as 3 (out of 10) category units.

Figure 10 shows that responses also assimilate toward the value of the prior response. This is expected since there is assimilation toward the prior stimulus, and stimuli and responses are correlated. This is consistent to the common finding in univariate data, that assimilation is greater to the prior stimulus when feedback is given and greater to the prior response when there is no feedback.

---insert Figure 10---

Another sequence effect is seen in Figure 11. This shows the

median time taken by the subjects to respond to the stimulus as a function of the difference between the current and prior stimulus. Responses were faster when successive stimuli were more similar, with similarity measured as the Euclidean distance between stimuli in the frequency-amplitude space. Response times and these stimulus differences are reliably correlated ($r = 0.87$, $p < 0.01$). Consistent with this, response times also correlate ($r = 0.68$, $p < 0.01$) with the difference between the current response and the prior response.

---insert Figure 11---

<u>No Feedback.</u> When feedback was not given, there was again assimilation between the response and the prior stimulus (Figure 12) and between the response and the prior response (Figure 13). Also again, the response times are correlated with the difference between successive stimuli ($r = 0.68$, $p < 0.01$) and with the difference between successive responses ($r = 0.72$, $p < 0.01$).

---insert Figure 12---

---insert Figure 13---

In AI data, responses are better correlated with the prior stimulus when feedback is given, but are better correlated with the prior response when feedback is not given (King & Lockhead, 1981). The similar result occurred with these sawtooth paired data. This pattern is consistent with a thesis of this research project. Performance is most associated with the best information available to the subjects. This best information is feedback when

feedback is given, and is the subjects' best estimate of that feedback (i.e., the prior response) when feedback is not given.

Discussion. There are systematic and stable SEs in these bivariate data. Analogous to univariate cases, responses assimilate toward the value of the prior stimulus and/or the value of the prior response. Also, the more different the successive stimuli, the more time is needed to respond.

Except that some subjects were more consistent than others, no individual differences were noted. For each subject and each condition, there are marked effects of sequence on choices and on response times. Choices are biased toward the response and stimulus of the prior trial, and it takes longer to respond when successive stimuli are more different.

Performance is often poorer in univariate than multivariate tasks. The results of this study indicate this difference is not directly attributable to SEs. SEs are essentially the same in both classes of data. The following study suggests the performance difference is related, instead, to how stimuli differ. Successive univariate stimuli can differ from one another in only one quantitative way. Multidimensional stimuli can differ quantitatively and qualitatively. If this is the important factor, then relations among attributes, not just individual attribute values are involved. In this case, the subjects might not judge only an attribute of interest to the experimenter, such as loudness or pitch, when stimuli vary in more than one way. Rather, as reported next, when people are asked to judge the loudness of a tone they appear to first process the entire stimulus and only then analyze and judge the attribute. If this

interpretation is correct, then whenever stimuli vary from trial to trial in something other than loudness, loudness judgments would be affected. The following study examined loudness judgments when stimuli varied from trial to trial in pitch.

<u>Classification by Loudness.</u> The preceding study showed tones composed of correlated amplitudes and frequencies are better identified than are univariate tones, and showed there are sequence effects in judgments of such bivariate stimuli. This study examined amplitude judgments when frequency varied randomly between trials. Two auditory amplitudes were paired orthogonally with two auditory frequencies, and people were asked to classify each tone according only to its loudness, i.e., to ignore its pitch.

This procedure is called a filtering task because the subjects' task is to ignore or "filter out" the irrelevant dimension (pitch in this case) and respond only in terms of the relevant dimension (loudness). This is also called an orthogonal task because the stimulus dimensions are combined orthogonally, such that information about the level of one dimension provides no information about the level of the other dimension.

When integral stimuli (Lockhead, 1966) are studied, performance is better (faster, fewer errors) if the stimuli do not vary from trial to trial on an irrelevant dimension (the univariate or control condition) than if they do vary between trials on an irrelevant dimension (the orthogonal condition). Accordingly, if loudness and pitch form integral stimuli, loudness judgments should be faster in the control (univariate)

condition than in the orthogonal condition. Furthermore, if the amount by which successive stimuli differ affects the magnitude of any SE or error due to shifted criterion placements, then there should be more interference when the irrelevant dimension varies more between trials. To examine this, various pitch differences were examined. Any effect on classification associated with the amount by which stimuli differ on an irrelevant dimension has not previously been reported for a filtering task.

Method and Procedure. Subjects judged loudness when pitch: (1) did not change between trials, (2) could differ a small amount between trials, and (3) could differ a large amount between trials.

The same auditory intensities were used in all conditions, 79 and 81 dB. There were two control (univariate) conditions. In one, the 79 and 81 dB amplitudes were always presented at 1000 Hz. In the other, these amplitudes were always presented at 1500 Hz. There were three orthogonal conditions, called narrow, intermediate, and wide range. The two amplitudes were presented randomly at 1000 and 1015 Hz (narrow range), or 1000 and 1045 Hz (intermediate range), or 1000 and 1500 Hz (wide range). The five conditions are diagrammed in Figure 14.

---insert Figure 14---

Six subjects gave 400 responses in each condition. They classified each tone as quiet or loud by pressing the left or right of two buttons.

Results and Discussion. The median response times across subjects for each condition are shown in Figure 15. For the following discussion, data represented by the solid and dashed lines are averaged at each pitch condition and considered as a single point. According to analysis of variance, the response times in the five conditions are reliably different [$F(4,8) = 4.82$, $p < 0.05$]. According to an orthogonal contrast analysis, performance is significantly fastest in the univariate tasks, slower when frequencies vary by 15 or 45 Hz, and still slower when the frequencies vary by 500 Hz. Error rates are positively correlated with response times, indicating speed-accuracy tradeoffs are not an important concern.

---insert Figure 15---

This observation that responses are slower in the orthogonal than the univariate condition replicates many earlier reports of integral stimuli (cf. Garner, 1974). A further observation here is that the magnitude of this effect is larger when the stimuli are more different on the irrelevant dimension. This is consistent with the general thesis of this project: Performance is more variable when successive stimuli are more different from one another. This is a range effect.

The magnitude of this range effect seems large. In comparison to the univariate conditions, classification by loudness was slowed by 100 msec. when the tones differed in frequency by 15 Hz, was slowed by 122 msec. when the tones differed by 45 Hz., and was slowed by 246 msec. when the tones

differed by 500 Hz.

Stimulus Range and Stimulus Sequence. In both AI and ME data, performance is poorer when the stimulus range is larger. This between-conditions effect is a range effect. Also in AI and ME data, performance on any trial is poorer when the prior stimulus is more different. This within-conditions effect is a sequence effect. Since successive stimuli are more different more often whenever the stimulus range is larger, range effects and sequence effects may be due to the same source.

If this is the case, then there should be sequence effects correlated with the range effects in these classification data. Performance should be poorer, within as well as between conditions, on trials that successive stimuli were more different.

To examine this, sequential response times were measured for each orthogonal condition. The results averaged over subjects are shown in Figure 15. The solid line shows the median response times on trials that the pitch repeated, i.e., the irrelevant dimension did not change between trials. The dashed line shows the median response times when the pitch level did change between trials; this never happens in univariate conditions.

Responses were about 60 msec slower when pitch changed between trials than when pitch repeated [$F\ (1,3) = 35.86$, $p < 0.01$], and these measures interact with the range of the irrelevant pitch dimension [$F = (2, 6) = 6.22$, $p > 0.05$].

Discussion. Trial-to-trial variation in pitch, an irrelevant dimension, affects performance when stimuli are classified according to loudness, the relevant dimension. Stimulus sequence also affects performance. Loudness judgments take more time and have more errors when pitch varies between trials than when pitch is constant from trial to trial. Further, the amount of this interference is greater when the pitch changes are greater.

These classification outcomes are consistent with AI and ME data. Whether people categorize, or identify, or judge a stimulus, the quality of performance decreases monotonically with increases in the difference between successive stimuli. At least for integral stimuli, this occurs whether or not that difference is along the dimension being judged or is along a nominally irrelevant dimension.

Conclusion

One goal of this research is to better understand what contextual factors affect judgments of sounds, and how those factors affect judgments. While the memory model in Figure 1 is intended to help examine such factors, the theme of this approach is that neither this nor any other simple model will be sufficient. This is because different information is available in different tasks and situations, and people use whatever information is available in order to maintain veridical response criteria.

Guesses as to what might have occurred in the absence of

stimuli or feedback begins an example. Here there is no information about the stimulus and people can only use internal strategies for selecting responses. Such guesses have SEs and these are associated with prior responses.

When feedback is given in such tasks, there are again large SEs. However, these are now associated more with the feedback than with prior responses.

If the stimulus generator is now turned on, so there are actual stimuli to be judged, subjects use that stimulus information. Then, there are SEs associated with prior stimuli. However, there are still SEs associated with prior responses and feedback. The presence of stimuli does not eliminate those. The reason is that successive memories of the stimuli assimilate toward one another. These resulting errors in memory must be adjusted by the response system in order to maintain a veridical response scale.

A specific example is useful. Consider when people are asked to judge the relative intensities of successive tones. When a stimulus repeats in such studies, this stimulus ratio ($S_N/S_{N-1} = 1$) is 1 and subjects should respond "1". They rarely do. Rather, the response tends to be greater than one if the prior stimulus (S_{N-2}) was more intense than S_{N-1} and S_N, and the response tends to be less than one if S_{N-2} was less intense than S_{N-1} and S_N. Although the average response is indeed about 1, this is simply an average; "1" itself is rarely given when the stimulus ratio is 1 (Lockhead & King, 1983).

This does not mean subjects avoid the response "1". They do not. Subjects often report "1" when the successive stimuli are different.

Both results are due to assimilation. S_{N-1} assimilates toward the memory of S_{N-2}. This means that M_{N-1} is different than S_{N-1}. S_N is compared with this memory. If S_{N-1} had assimilated to a smaller value, then the response will be greater than 1 when the stimulus repeats; if assimilation had been to a larger value, then the response will be less than 1 when the stimulus repeats.

This effect is not trivial. Successive-ratio judgments have been affected by as much as a factor of 8 when loudnesses were judged (Lockhead & King, 1983).

Multidimensional Stimuli

This conclusion section has so far considered only univariate stimulus sets. These are commonly difficult to identify precisely. Multivariate stimuli are often more readily identified. To account for this difference, we considered the possibility that multidimensional stimuli are easy to identify because the large SEs in univariate judgments are somehow reduced or eliminated in bivariate tasks. It was not known if this is the case because SEs had not previously been reported when multidimensional stimuli were identified.

The consideration is wrong. There is marked assimilation when sawtooth-paired tones are identified. Thus, multidimensional tones are not identified easily because SEs are reduced. Another

interpretation is needed.

One relevant issue is whether or not attributes of tones in multidimensional sets are judged independently. To examine this, we had people classify tones according to loudness when the tones also varied, or did not vary, in pitch from trial to trial. These pitch variations had a marked interference effect on loudness classification. This indicates the tones are integral stimuli. In addition, the magnitude of this interference effect was greater when pitch varied more between trials (cf. Figure 15).

We interpret this interference as meaning that subjects initially locate the integral stimulus in some memory or perceptual space, and then select a response in terms, jointly, of this perceived location and known possible locations (the IDF).

The sawtooth-identification data are consistent with this interpretation. Sawtooth-paired stimuli are separated further from one another in the two-dimensional space than are the comparable univariate stimuli. While there is assimilation in both instances, univariate and bivariate tasks, these redundant stimuli are sufficiently separated that less precise identification is needed to identify them than to identify the comparable univariate stimuli.

Such a spatial metaphor is not necessary. The situation can also be described propositionally. For example, if the subject judges the loudness of a univariate stimulus to be low, say it is judged to be #1, #2, or #3, then any of these 3 responses might

be selected for the loudness value. But if the stimulus is sawtooth-paired with pitch, and if pitch is also identified only approximately, then these three responses are not equally likely candidates. If the actual loudness had been #1, then the pitch was low (#3); if the loudness had been #2, then the pitch was intermediate (#6); and if the loudness had been #3, then the pitch was high (#9). By concatenating the two judgments, #1, or #2, or #3 will be selected. Although loudness and pitch are each identified only approximately (low, medium, or high for each), the stimulus can be identified precisely.

In summary, for univariate and multivariate stimulus sets, whether examined in ME, AI, or classification tasks, there are always extensive SEs, the memories of prior stimuli are biased, the subject rarely knows the precise value of any attribute of a stimulus, and subjects use whatever information is available to maintain a veridical response scale.

References.

Baird, J. C. & Noma, E. (1975) Psychological study of numbers: I. Generation of numerical responses. Psychological Research, 37, 281-297.

Garner, W. R. (1974) The Processing of Information and Structure, Erlbaum, N.J.

Green, D., Luce, R. & Duncan, J. (1977) Variability and sequential effects in magnitude production and estimation of auditory intensity. Perception & Psychophysics, 22, 450-456.

Green, D. M. & Swets, J.A. (1974) Signal Detection Theory and Psychophysics. NY: R. E. Krieger Pub. Co.

Holland, M. & Lockhead, G. (1968). Sequential effects in absolute judgments of loudness, Perception & Psychophysics, 3, 409-414.

Hinson, J. & Lockhead, G.R. (1986) Range effects in successive discrimination procedures. Journal of Experimental Psychology: Animal Behavior Processes, 12, 270-276.

Hinson, J. & Lockhead, G.R. (1987) Momentary and Global Maximizing, Behavioral Processes, 14, 21-34.

Holland, M. (1968) Channel capacity and sequential effects: The

influence of the immediate stimulus history on recognition performance. Ph.D. dissertation, Duke Univ.

Holland, M, & Lockhead, G. (1968) Sequential effects in absolute judgments of loudness. *Perception & Psychophysics*, *3*, 409-414.

Jesteadt, W., Luce, D., & Green, D. (1977) Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, *3*, 92-104.

King, M.C. (1980) Sequential effects in the variability of magnitude estimates of loudness. Ph.D. dissertation, Duke University.

King, M.C., Gruenewald, P., & Lockhead, G. R. (1978) Holistic vs. attribute repetition effects in classifying stimuli. *Memory and Cognition*, *6*, 438-445.

King, M. C. & Lockhead, G. (1981) Response scales and sequential effects in judgment, *Perception & Psychophysics*, *30*, 599-603.

Lockhead, G. R. (1966) Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, *72*, 95-104.

Lockhead, G. R. (1984) Sequential predictors of choice. In, *Preparatory States and Processes*, S. Kornblum and J. Requin (Eds.), Erlbaum, 27-48.

Lockhead, G. & Hinson, J. (1986) Range and Sequence Effects in Judgment, Perception & Psychophysics, 40, 53-61.

Lockhead, G. & King, M. C. (1983) A memory model of sequential effects in scaling tasks, Journal of Experimental Psychology: Human Perception and Performance, 9, 461-473.

Lockhead, G. R. & King, M. C. (1986) Variance in magnitude estimations, Fechner Day '86, International Society of Psychophysics, Stockholm, 59-64.

Lockhead, G. R. & King, M. C. (1988) An alternative to the Neural-Attention Hypothesis for Auditory Psychophysics, Fechner Day '88, International Society of Psychophysics, Stirling, England, 101-102.

Luce, R. D. & Green, D. M. (1978) Two tests of a neural attention hypothesis for auditory psychophysics. Perception & Psychophysics, 23, 363-371.

Luce, R. D. & Krumhansl, C. L. (1988) Measurement, Scaling, and Psychophysics. In R. Atkinson, R. Herrnstein, G. Lindzey, & R. D. Luce Stevens' Handbook of Experimental Psychology (2nd Ed.), NY, Wiley Interscience.

Staddon, J., King, M. & Lockhead, G.R. (1977) Sequential effects in absolute judgments. Journal of Experimental Psychology: Human Perception and Performance, 6, 290-301. Stevens, S.S. (1975) Psychophysics: Introduction to its perceptual, neural, and social prospects. G. Stevens (Ed.), NY, Wiley.

Triesman, M. & Williams (1984) T. A theory of criterion setting with an application to sequential dependencies. Psychological Review, 91, 68-111.

Ward, L. & Lockhead, G.R. (1970) Sequential effects and memory in category judgments. Journal of Experimental Psychology, 1970, 84, 27-34.

Ward, L. & Lockhead, G.R. (1971) Response system processes in absolute judgment. Perception & Psychophysics, 9, 73-78.

Figure Legends

<u>Figure 1.</u> The stimulus, S_N, is assumed to assimilate toward the memory of the prior stimulus, S_{M-1}, and is thus overestimated in the response, R_N. (From Lockhead & King, 1983, Figure 4).

<u>Figure 2.</u> Nine sensation functions reported by Stevens (1975, p. 113).

<u>Figure 3.</u> Average magnitude estimations given to thirty tones by subjects S.L and E.M.

<u>Figure 4.</u> Response frequencies (cell entries) to each stimulus by each subject in the ME task summarized in Figure 3.

<u>Figure 5.</u> The average response error on trial N when the feedback on trial N + k (the abscissa) was small (filled symbols) or large (open symbols). [From Ward & Lockhead, 1971.]

<u>Figure 6.</u> The correlation coefficients between successive responses as a function of the intensity difference between the current and previous stimulus for ME data averaged over two subjects.

<u>Figure 7.</u> The correlation coefficients between successive responses as a function of the absolute intensity difference between the current and previous stimulus for ME-with-feedback data averaged over subjects.

<u>Figure 8.</u> Linear correlated (open circles) and sawtooth correlated (filled circles) pitch-loudness dimensions.

<u>Figure 9.</u> Average responses to sawtooth correlated stimuli (cf. Figure 10) when feedback was given, as a function of the prior stimulus. Enclosed regions indicate the 10 stimuli. Numerals indicate the intensity of the prior stimulus. Positions of the numerals show the median response to each stimulus, in X-Y co-ordinates in terms of Figure 10, as when the indicated stimulus occurred on the prior trial.

<u>Figure 10.</u> Average responses to sawtooth correlated stimuli (cf. Figure 10) when feedback was given, as a function of the prior response.

<u>Figure 11.</u> Median response times to identify each sawtooth-paired stimulus as a function of the difference between it and the prior stimulus when feedback was given.

<u>Figure 12.</u> Average responses to sawtooth correlated stimuli (cf. Figure 10) when feedback was not given, as a function of the prior stimulus. Enclosed regions indicate the 10 stimuli.

Numerals indicate the intensity of the prior stimulus. Positions of the numerals show the median response to each stimulus, in X-Y co-ordinates in terms of Figure 8, as when the indicated stimulus occurred on the prior trial.

Figure 13. Average responses to sawtooth correlated stimuli (cf. Figure 10) when feedback was not given, as a function of the prior response. Identical to Figure 12, except numerals now indicate the prior response.

Figure 14. Auditory frequencies when subjects classified sinewaves according to loudness in the five sorting tasks Q = 79 dB; L = 81 dB).

Figure 15. Median response times to classify loudnesses when pitch repeated (solid lone) and when pitch changed (dashed line) between trials.

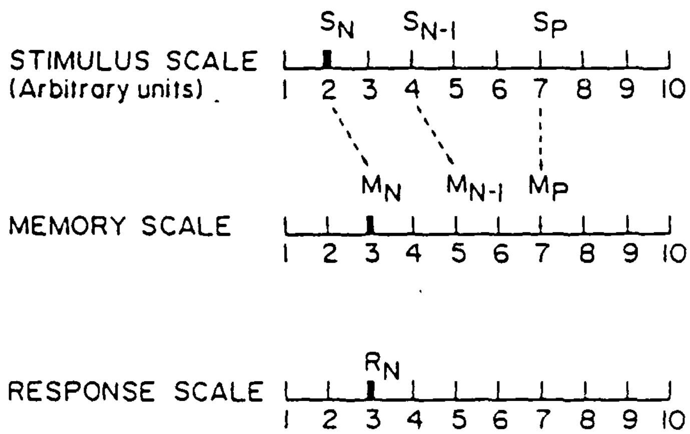


Figure 1. The stimulus, S_N, is assumed to assimilate toward the memory of the prior stimulus, S_{M-1}, and, is thus overestimated in the response, R_N, for this example.

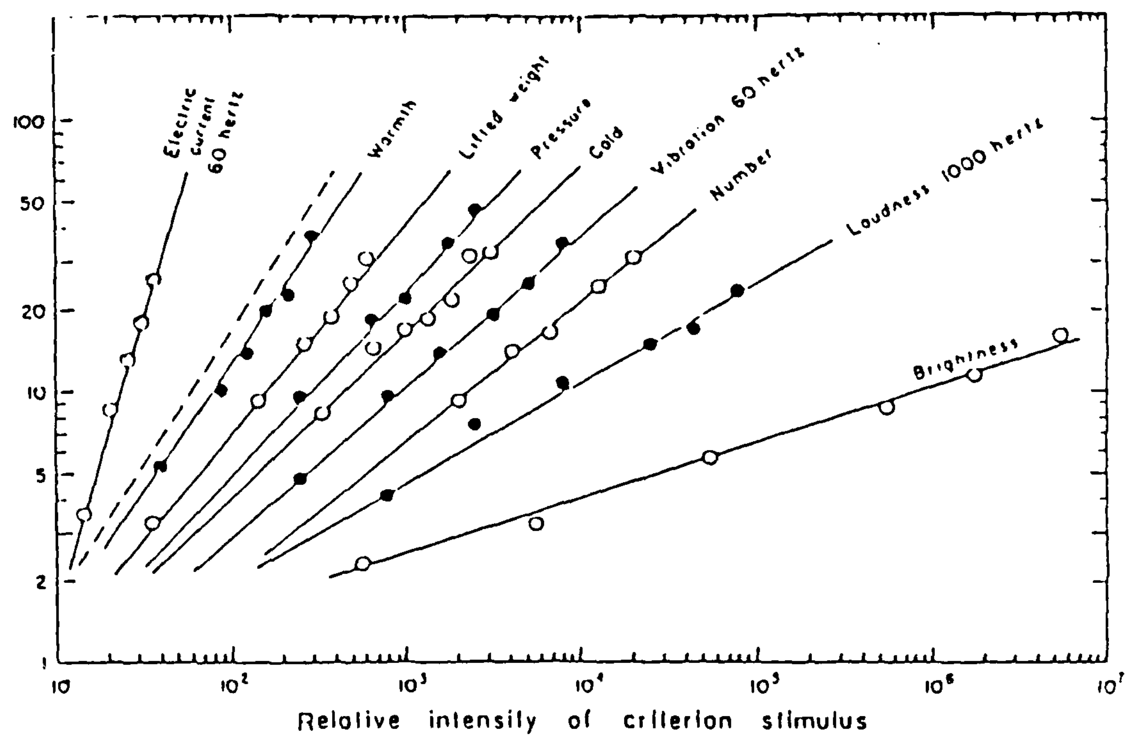


Figure 2. Nine sensation functions. (From Stevens, 1975, p. 113).

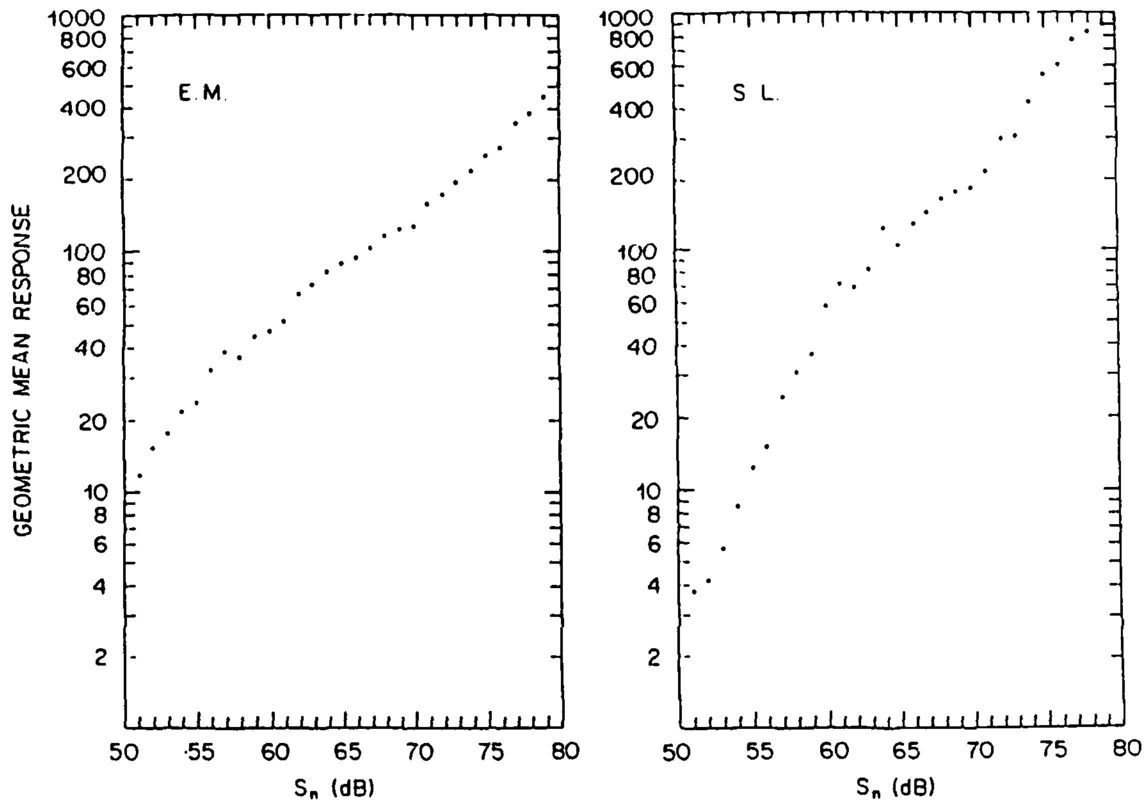


Figure 3. Average magnitude estimations of thirty tones by subjects S.L and E.M.

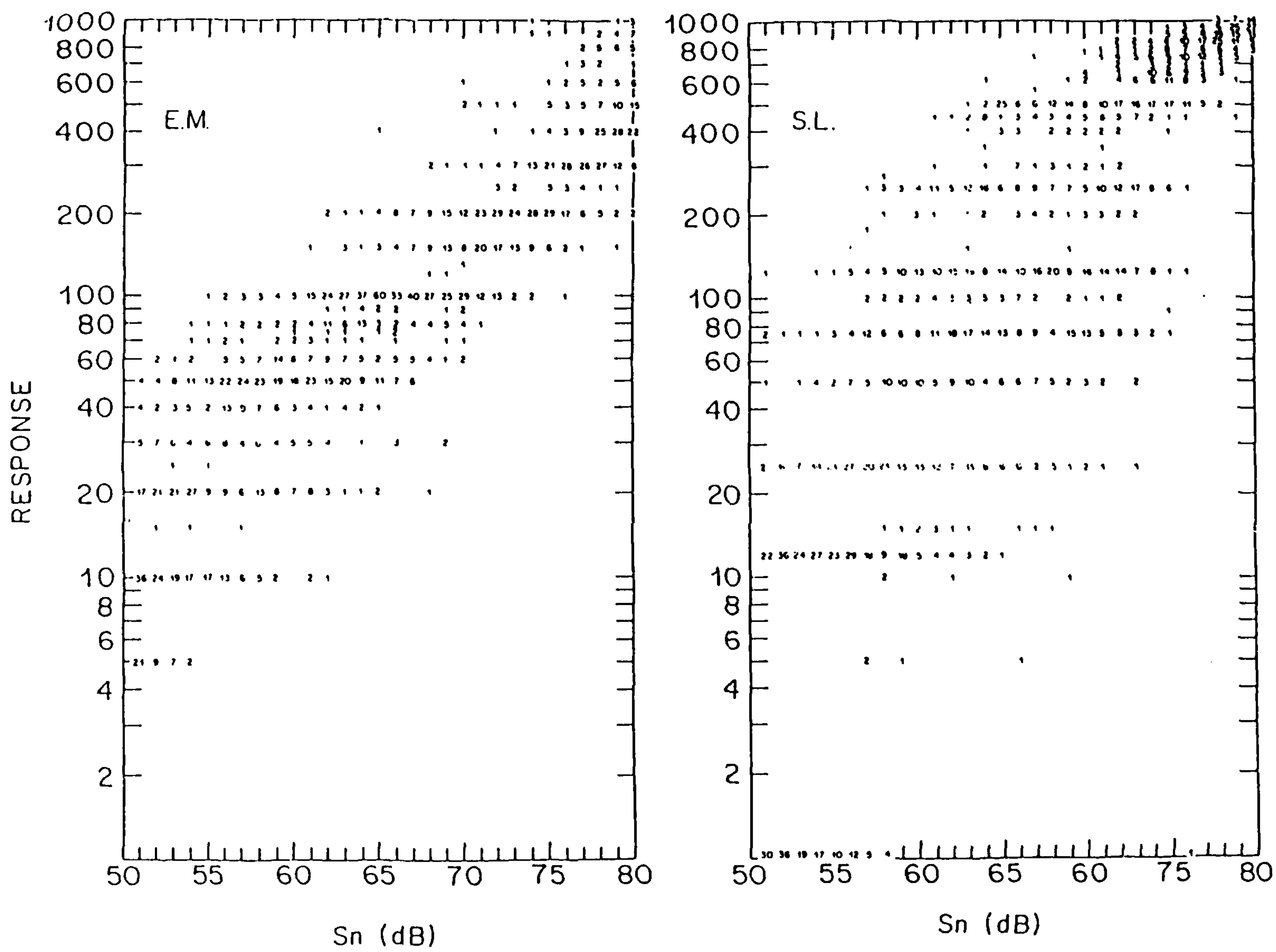


Figure 4. Frequencies (cell entries) of each response to each stimulus by S.L. and E.M. These data were averaged to produce Figure 3.

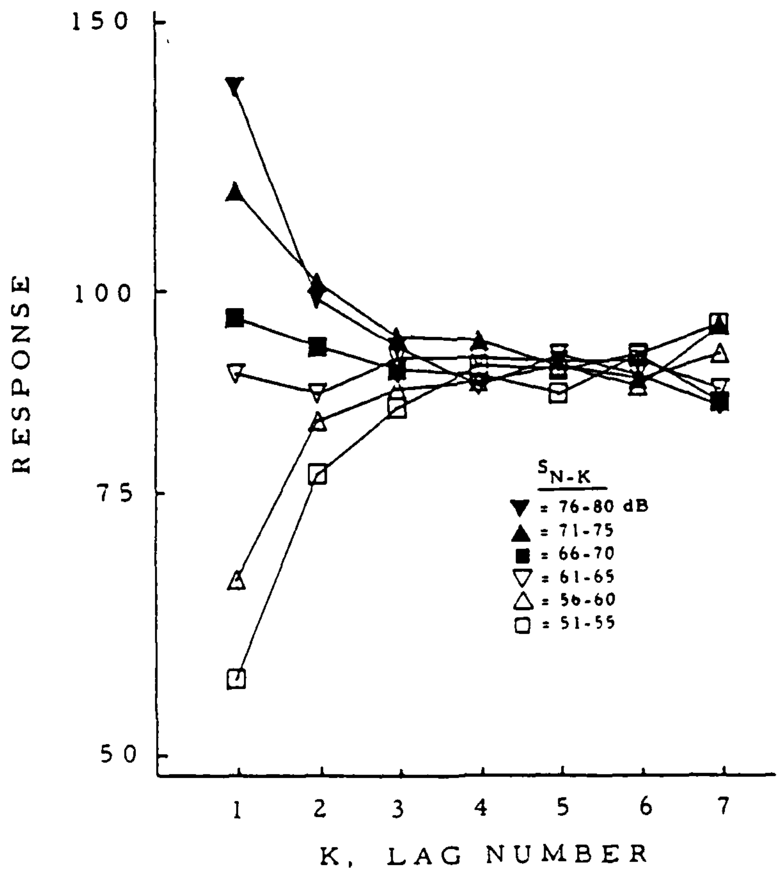


Figure 5. The average response on trial N when the feedback on trial N - k was small (filled symbols) or large (open symbols).

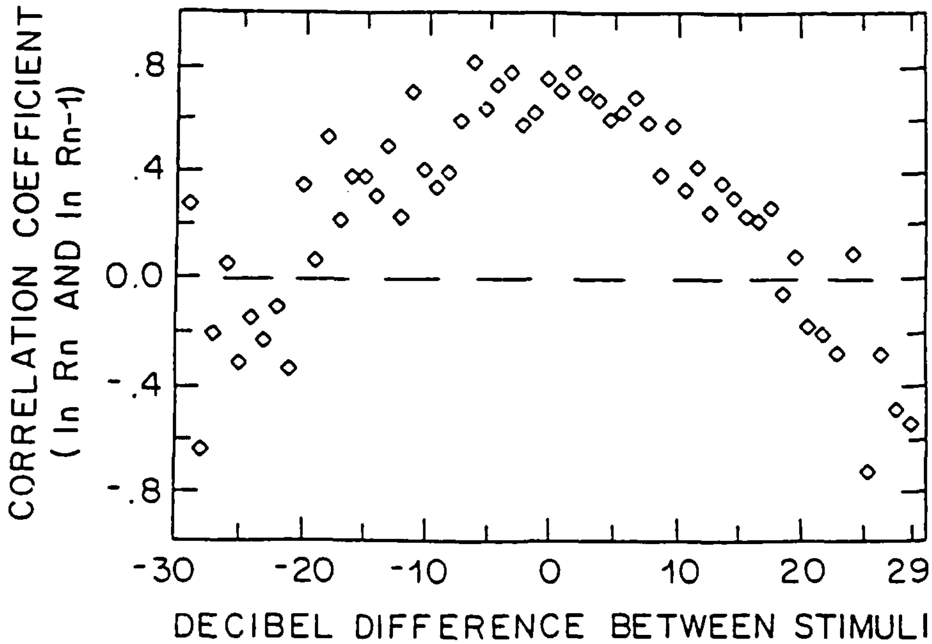


Figure 6. Correlations between successive responses as a function of the intensity difference between successive stimuli for ME data combined over subjects.

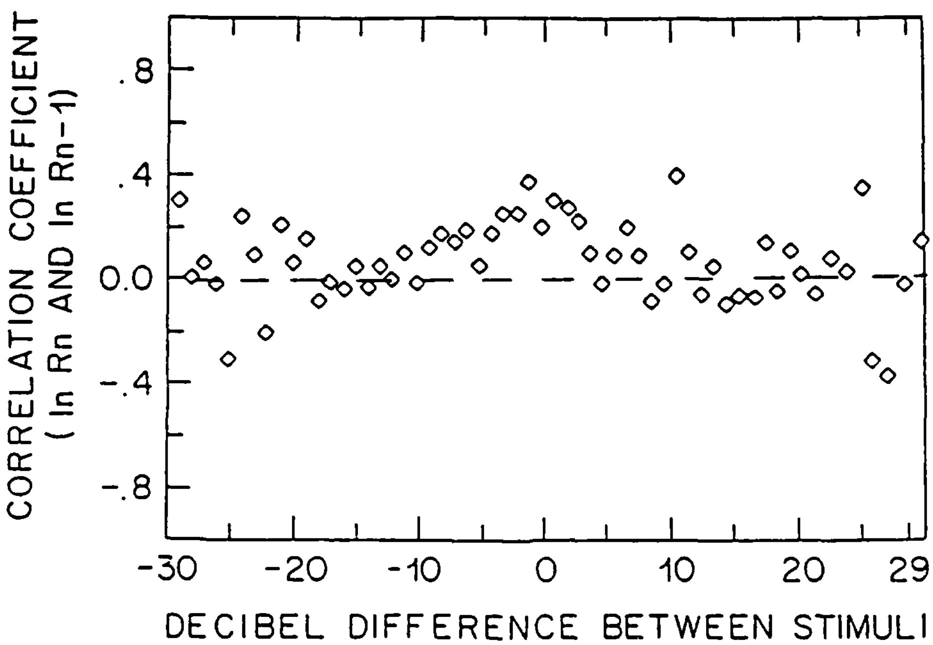


Figure 7. Correlations between successive responses as a function of the intensity difference between successive stimuli for ME-with-feedback data combined over subjects.

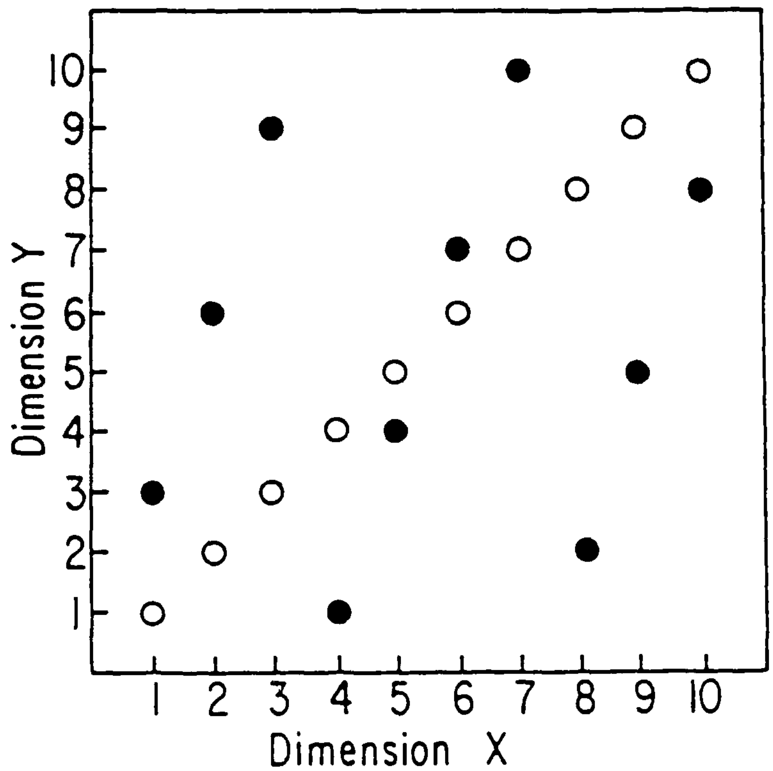


Figure 8. Linear correlated (open circles) and sawtooth correlated (filled circles) pitch-loudness stimulus sets.

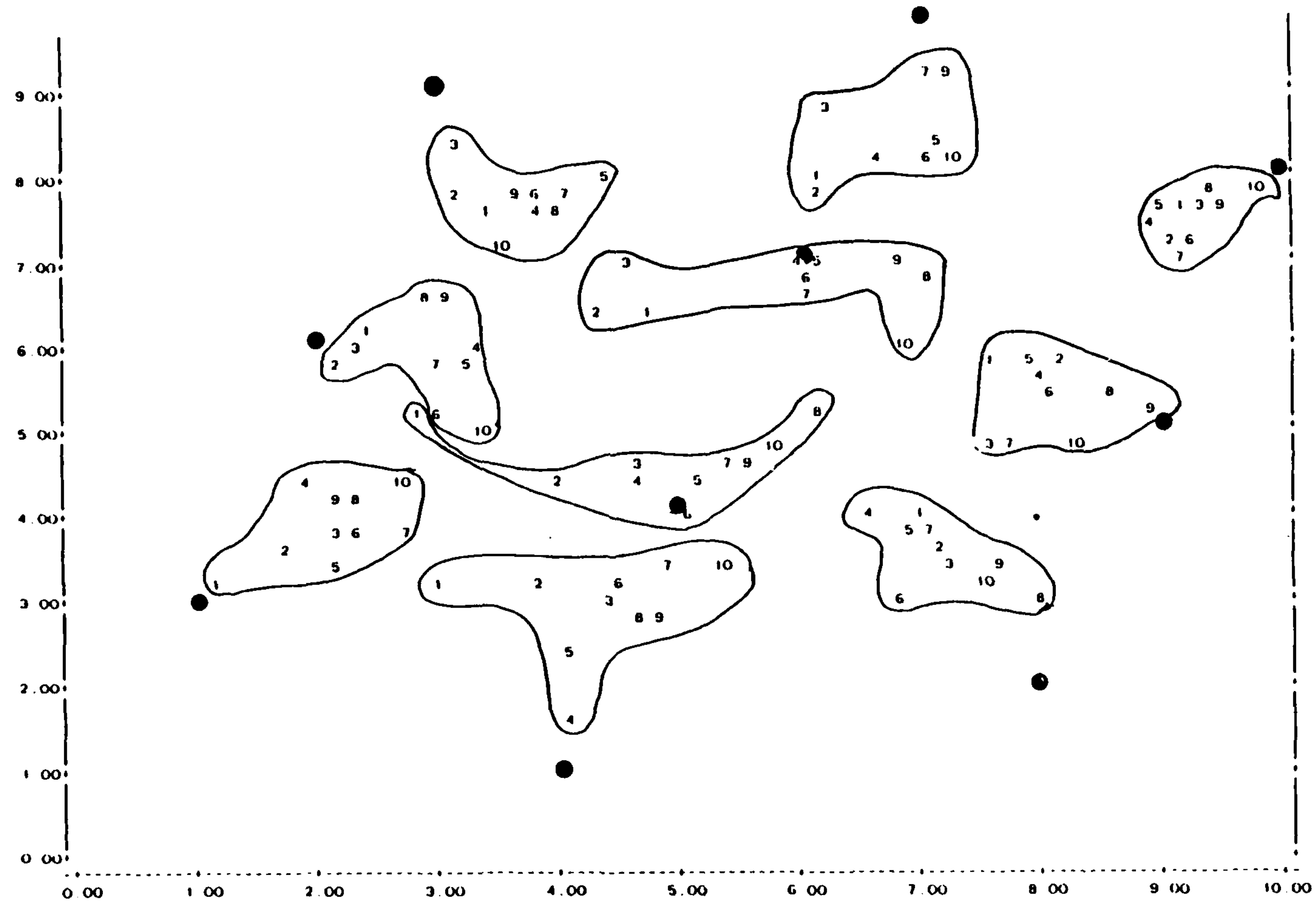


Figure 9. Average responses to sawtooth-paired stimuli when feedback was given, as a function of the prior stimulus. Outlined regions indicate the stimulus on the current trial (cf. solid dots in Figure 8). Numerals within each outline indicate the stimulus on the prior trial. The positions of the numerals show the median response to that stimulus and that prior stimulus in X-Y coordinates.

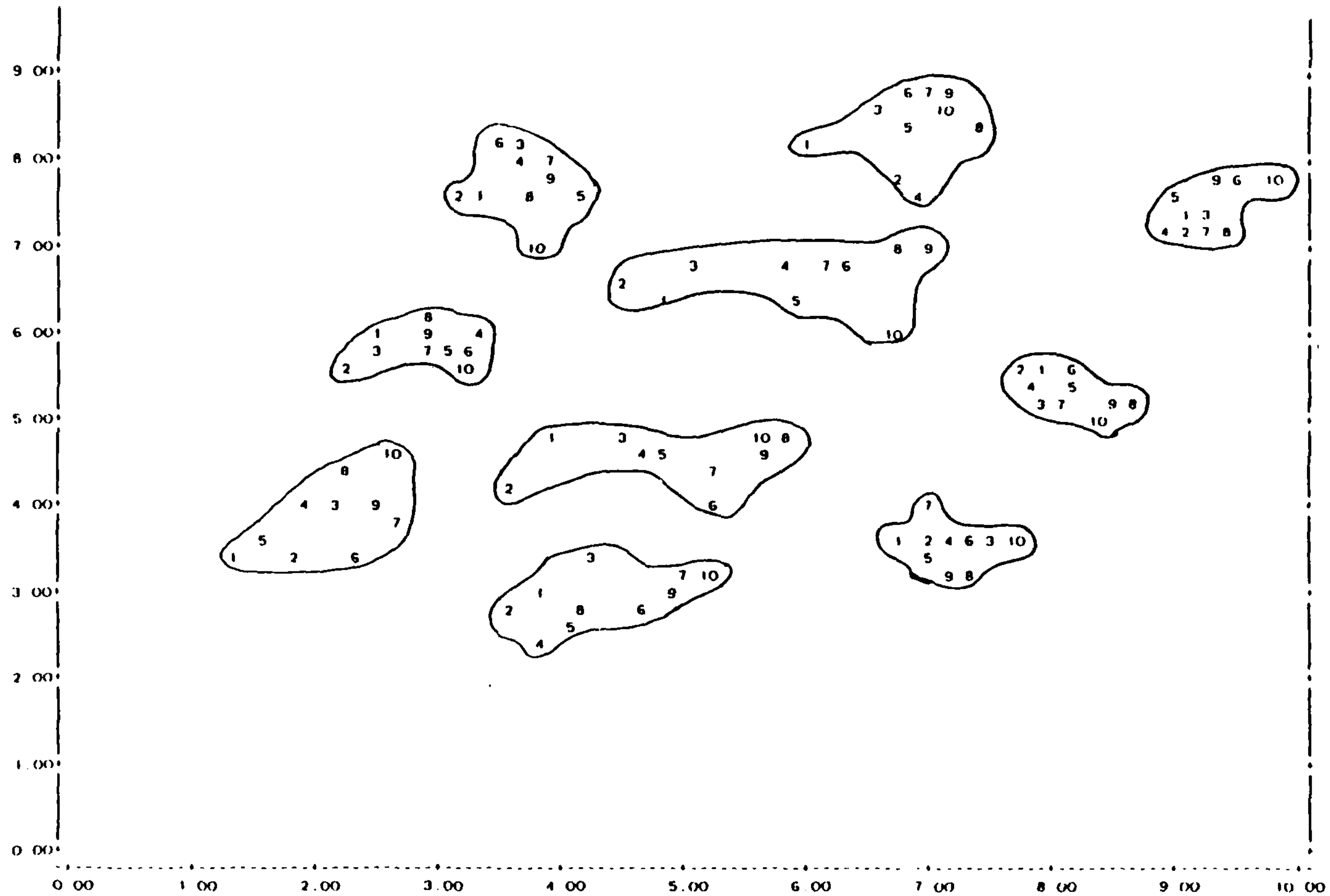


Figure 10. Average responses to sawtooth correlated stimuli when feedback was given, as a function of the prior response. Identical to Figure 9, except numerals indicate the prior response rather than the prior stimulus.

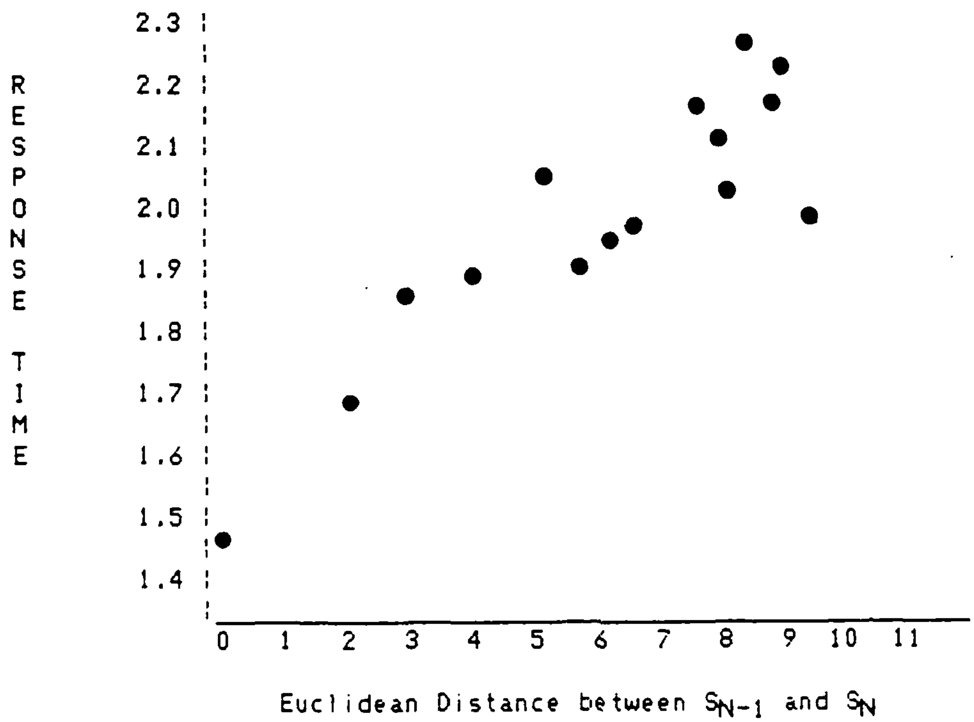


Figure 11. Median response times to identify each sawtooth-paired stimulus as a function of the distance, in the X-Y co-ordinates of Figure 8, between it and the prior stimulus, when feedback was given.

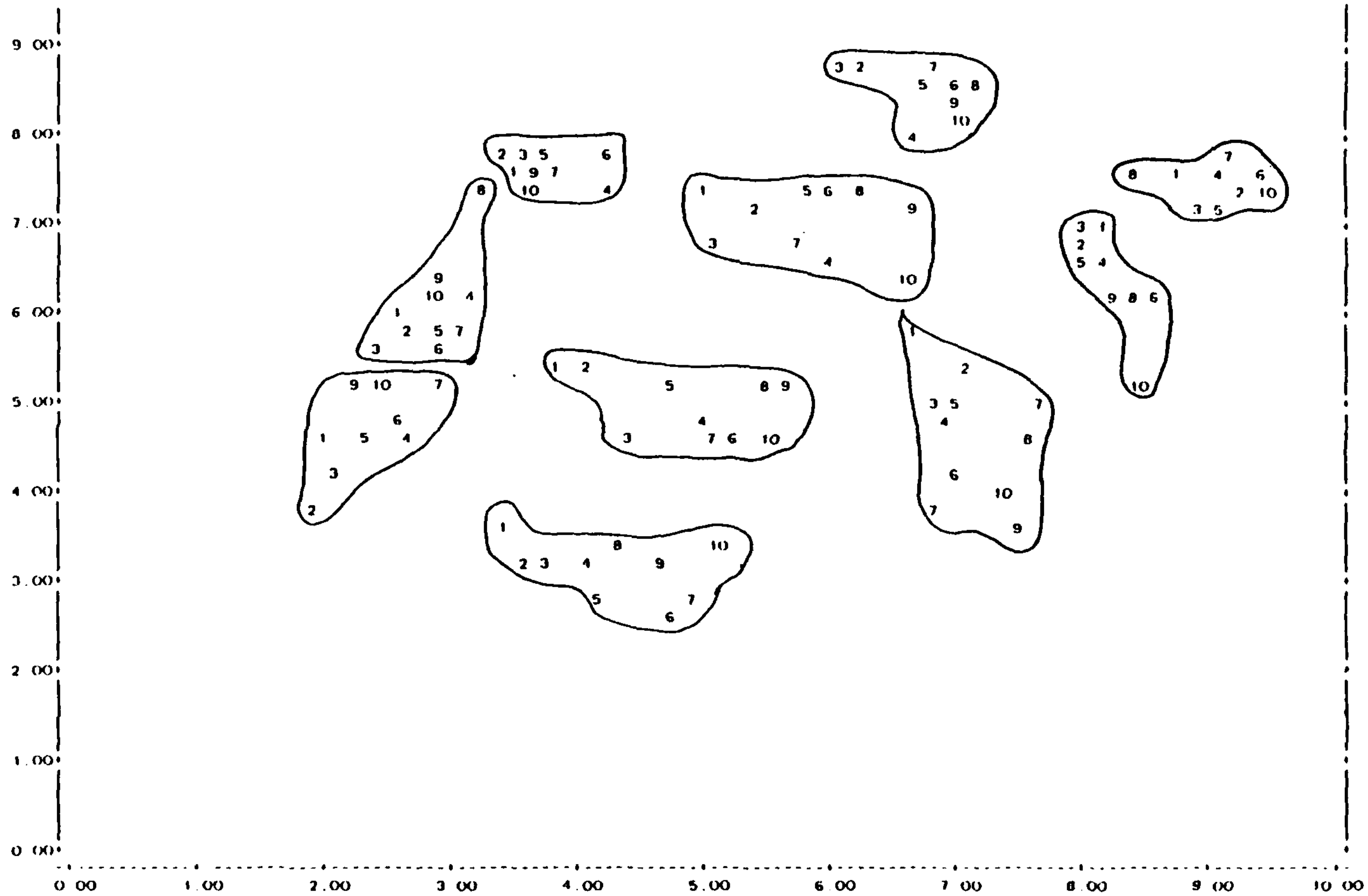
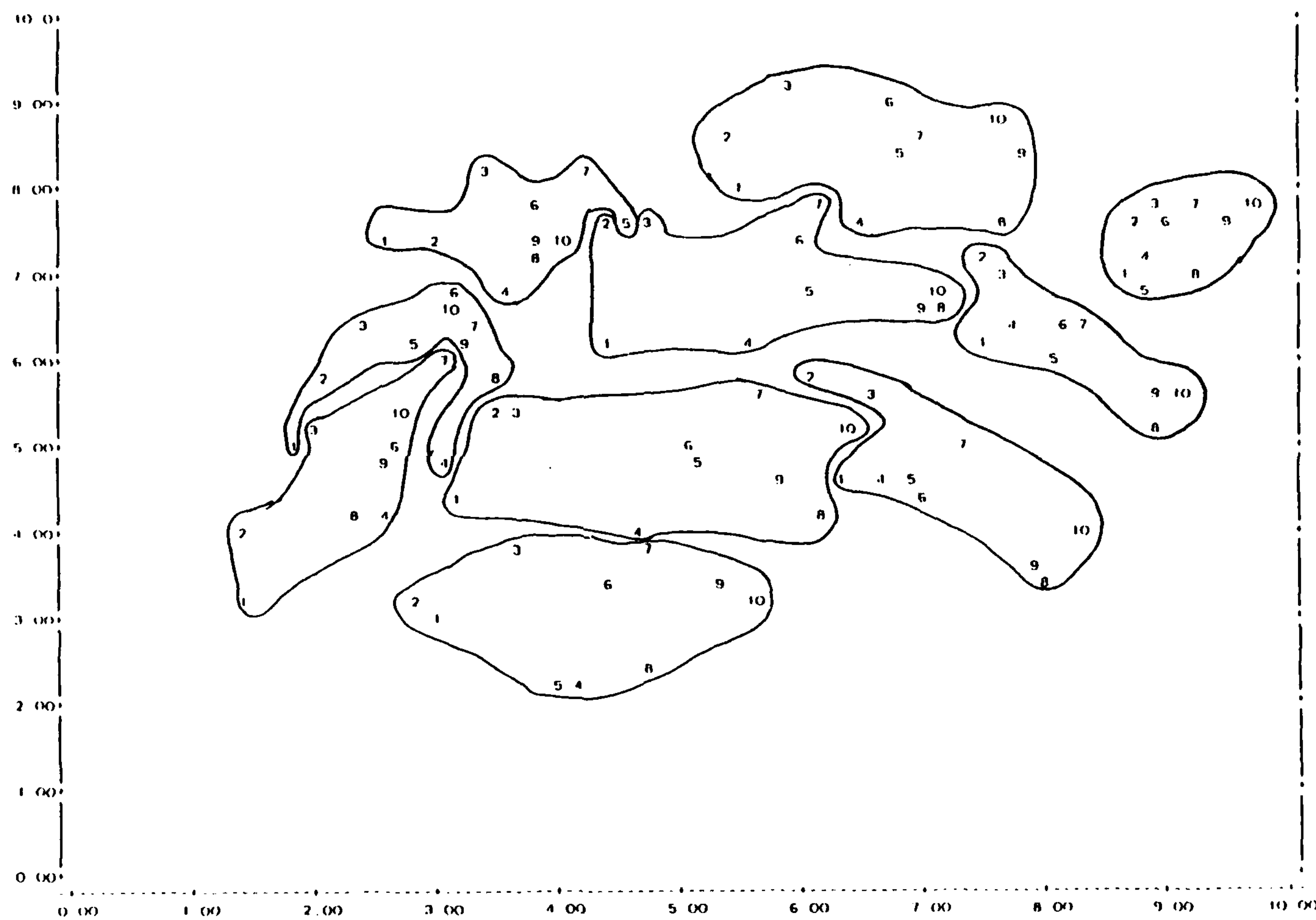


Figure 12. Average responses to sawtooth-paired stimuli when feedback was not given, as a function of the prior stimulus. Identical to Figure 9 except no feedback was given.



<u>Figure 13.</u> Average responses to sawtooth-paired stimuli when feedback was not given, as a function of the prior response. Identical to Figure 12, except numerals indicate the prior response rather than the prior stimulus.

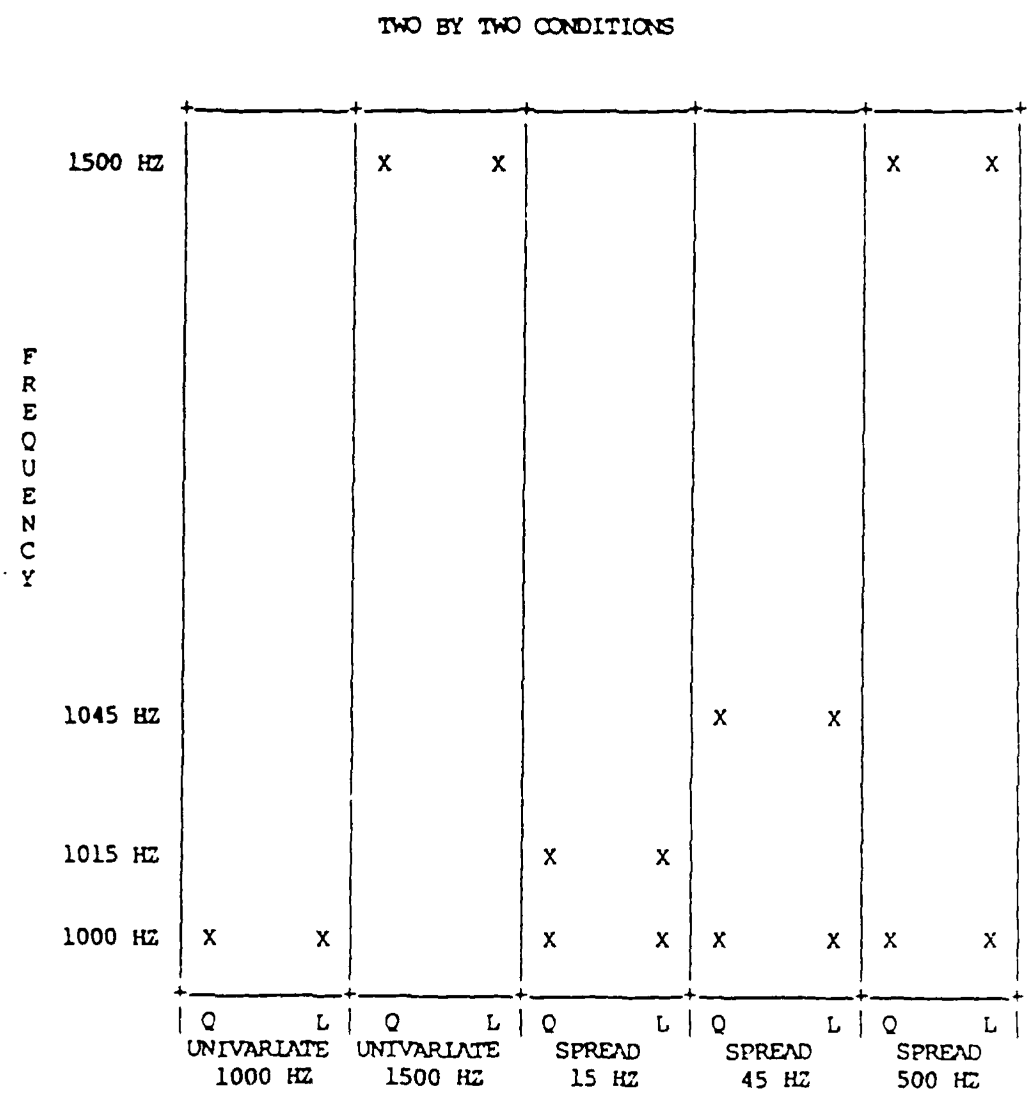


Figure 14. Auditory frequencies when subjects classified sinewaves according to loudness in the five sorting tasks (Q = 79 dB; L = 81 dB).

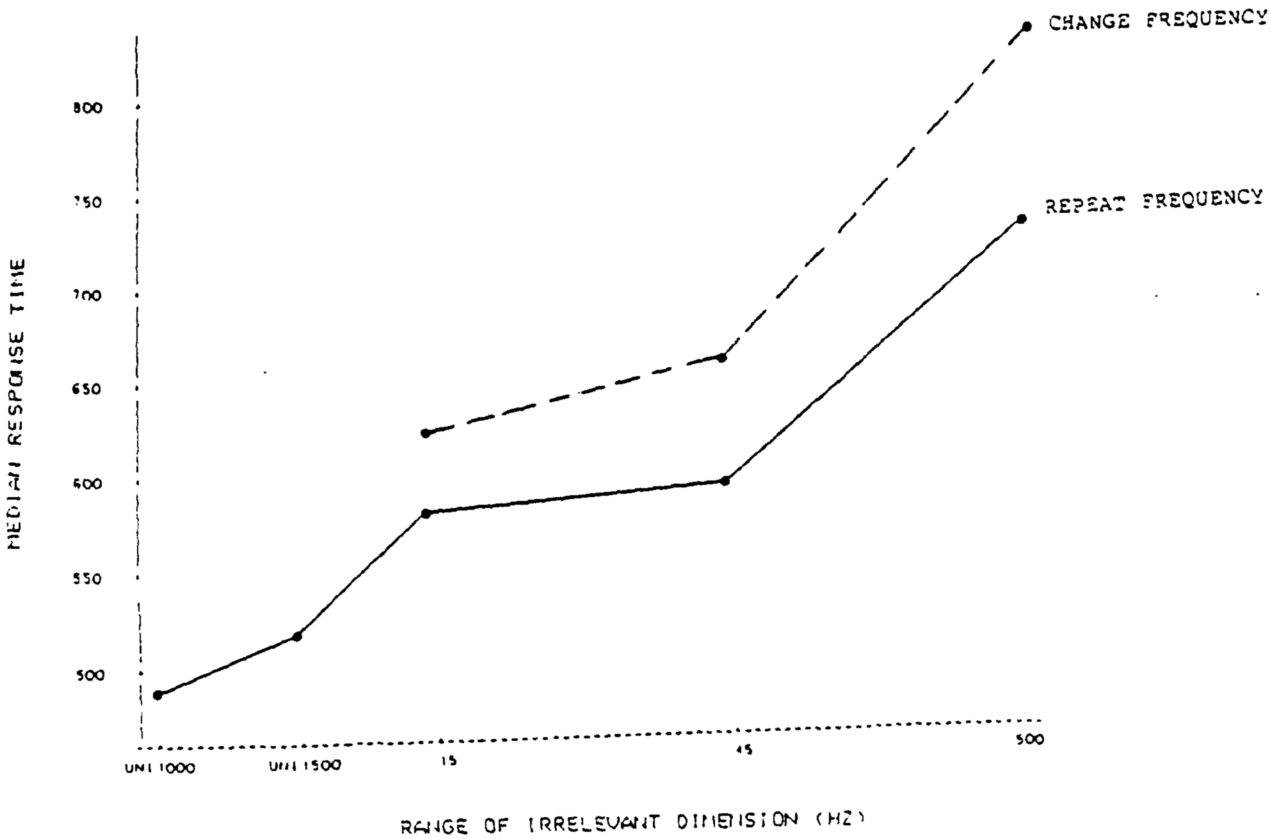


Figure 15. Median response times to classify loudnesses when pitch repeated (solid line) and when pitch changed (dashed line) between trials.